ELSEVIER

Theoretical Computer Science 281 (2002) 219–233

Theoretical Computer Science

www.elsevier.com/locate/tcs

Réduction de la non-linéarité des morphismes d'arbres Recognizable tree-languages and non-linear morphisms

Max Dauchet*, Sophie Tison, Marc Tommasi

LIFL, UPRESA 8022 CNRS UA 369, Bâtiment M3, Université de Lille I, 59655 Villeneuve d'Ascq Cedex, France

Abstract

We try to understand certain complex phenomena which appear in the study of tree morphisms in the non-linear case. We prove that every transformation performed by a quasi-alphabetic homorphism that maps a recognizable tree language to another one can be done "without using non-linearity". As a corollary, we get an extension of the cross-section theorem.

Résumé

Le but de ce papier est d'étudier certaines propriétés des morphismes d'arbres et d'éclairer certains phénomènes complexes propres aux arbres. En particulier, on montre que si l'image par un morphisme quasi-alphabétique d'une forêt reconnaissable est reconnaissable, on peut ≪réduire la non linéarité≫ du morphisme (en un sens que l'on précisera). Ce résultat nous permet de généraliser le théorème de cross-section.

1. Introduction

Dans ce papier nous étudions certaines propriétés propres aux morphismes d'arbres. En effet, si les principaux concepts—telle la reconnaissabilité par automate fini—définis sur les forêts sont des généralisations naturelles du cas des langages de mots, de nombreuses propriétés ne se transposent pas: en particulier, un phénomène nouveau et important est celui de la *non-linéarité*. Un terme t est dit non linéaire si il contient plusieurs occurrences d'une même variable. La manipulation de termes non linéaires est courante en programmation logique, fonctionnelle ou équationnelle mais

* Corresponding author.

E-mail addresses: max.dauchet@lifl.fr (M. Dauchet), tison@lifl.fr (S. Tison), tommasi@lifl.fr (M. Tommasi).

0304-3975/02/$ - see front matter
PII: S0304-3975(02)00014-2

les ≪outils≫ à notre disposition sont peu nombreux. Par exemple, l'ensemble des instances closes d'un terme non-linéaire n'est pas reconnaissable et donc l'ensemble des formes réductibles d'un système de réécriture n'est pas nécessairement reconnaissable si le système n'est pas linéaire. Ou encore, pour citer un autre exemple qui marque la différence entre le cas linéaire et le cas général dans le domaine de la réécriture, un système de réécriture sans paire critique est confluent dans le cas linéaire gauche, non confluent dans le cas général.

Face à cette situation, on peut envisager plusieurs approches; une première approche consiste à définir de nouveaux outils qui permettent de manipuler des ≪égalités≫, voire des différences; plusieurs formalismes ont été proposés, comme les automates à tests [7,4,6,5] ou les grammaires synchronisées [14]. Une démarche complémentaire consiste à réduire la non-linéarité quand c'est possible. Le principe général de cette démarche peut s'énoncer grossièrement par: ≪quand la reconnaissabilité est conservée, la non-linéarité peut être évitée (dans un sens à préciser)≫. Par exemple, Kucherov et Tajine, Hofbauer et Huber, Gilleron et Vàgvölgyi étudient les systèmes de réécriture dont l'ensemble des formes réductibles est reconnaissable et les transforment en systèmes linéaires [11,12,13,17].

Dans cet article, nous appliquons cette démarche au cas des morphismes. En effet, l'image homomorphe d'une forêt reconnaissable est reconnaissable si le morphisme est linéaire, non-nécessairement reconnaissable sinon—et donc la famille des langages reconnaissables d'arbres n'est pas close par morphisme, contrairement au cas des mots. Cependant, il existe des images par morphisme non-linéaire de forêts reconnaissables—même infinies—qui sont reconnaissables. C'est surtout cette situation que nous allons examiner ici. L'idée intuitive est la suivante: si une forêt reconnaissable contient une infinité d'arbres de la forme $t(v_i, v_i)$, elle contient aussi des arbres de la forme $t(v_i, v_j)$, avec v_i différent de v_j; en exploitant ceci, on peut montrer le résultat de base suivant:

Théorème. *Soit Φ un morphisme quasi-alphabétique et K une forêt reconnaissable; si $\Phi(K)$ est reconnaissable, il existe une forêt reconnaissable R et un morphisme Π linéaire, quasi-alphabétique, complet et strict tel que $\Phi \circ \Pi$ soit linéaire, quasi-alphabétique, complet et strict, $\Pi(R) \subset K$ et $\Phi \circ \Pi(R) = \Phi(K)$. (Les définitions des différentes classes suivent l'introduction.)*

Ce résultat technique permet d'étendre le théorème de transversale rationnelle (ou ≪cross-section≫): soit R une forêt reconnaissable, Φ un morphisme, le théorème établit l'existence d'une forêt reconnaissable R' incluse dans R telle que Φ soit une bijection entre R' et $\Phi(R)$. Le résultat a été étendu des mots aux arbres quand Φ est linéaire [1]; il est en général faux pour un morphisme d'arbres quelconque. Nous conjecturons que l'hypothèse ≪Φ linéaire≫ peut être affaiblie en ≪$\Phi(R)$ reconnaissable≫; nous prouvons ici cette conjecture dans le cas où Φ est quasi-alphabétique.

On peut appliquer aussi ce résultat aux bimorphismes et en particulier, aux bijections rationnelles [15]; on peut ainsi énoncer:

Théorème. *Toute bijection entre forêts reconnaissables réalisée par un bimorphisme quasi-alphabétique peut être réalisée par un bimorphisme linéaire.*

Nous conjecturons que l'hypothèse ≪quasi-alphabétique≫ est superflue.

2. Notations et définitions

Soit Σ un alphabet gradué et $X=\{x_1,\ldots,x_i,\ldots\}$ un ensemble dénombrable de variables. Nous notons X_n l'ensemble fini des n variables $\{x_1,\ldots,x_n\}$. L'ensemble des termes sur Σ indexés par X (resp. par X_n) est noté $T(\Sigma,X)$ (resp. $T(\Sigma,X_n)$) et $T(\Sigma)$ est l'ensemble des termes clos sur Σ, i.e. sans variable. Un p-uple de termes est noté $\vec{v}=(v_1,\ldots,v_p)$ et l'ensemble des p-uples de termes de $T(\Sigma,X_n)$ est $T(\Sigma,X_n)^p$. L'ensemble des termes dont les occurrences de variables lues de gauche à droite sont exactement $x_1,\ldots,x_n$ est désigné par $CT_n(\Sigma)$. Les termes de $CT_n(\Sigma)$ sont dits d'arité n. Nous désignons par Σ_n l'ensemble des lettres de Σ d'arité n. Tout a de Σ_n s'identifie alors à $a(x_1,\ldots,x_n)$ et $x_1,\ldots,x_n$ sont appelées variables de a. L'ensemble des termes contenant exactement une occurrence de chaque variable dans X_n est noté $\widehat{CT}_n(\Sigma)$.

Soient Σ et Δ deux alphabets gradués. Un morphisme Φ de $T(\Sigma,X)$ dans $T(\Delta,X)$ est défini par une application de Σ dans $T(\Delta,X)$. Le morphisme Φ est *linéaire* (resp. *complet*) [1] si pour tout a de Σ, chaque variable de a apparaît au plus une fois dans $\Phi(a)$ (resp. si chaque variable de a apparaît au moins une fois dans $\Phi(a)$). Φ est *strict* si l'image d'une lettre contient au moins une lettre. Une *quasi-lettre* est un terme de la forme $a(v_1,\ldots,v_n)$ où a est une lettre et chaque v_i est soit un terme clos, soit une variable. Un morphisme est *alphabétique* (resp. *quasi-alphabétique*) si l'image d'une lettre est une lettre (resp. une quasi-lettre) ou une variable.

Soit t un arbre de $T(\Sigma,X)$; on note $|t|$ sa *hauteur*, i.e. le nombre d'occurrences de lettres sur un plus long chemin. pour un morphisme Φ, $|\Phi|=\max_{\alpha\in\Sigma}\{|\Phi(\alpha)|\}$.

L'ensemble des variables qui apparaissent dans t est $Var(t)$. En particulier, si Φ est un morphisme et a une lettre, $Var(\Phi(a))$ sera confondu avec l'ensemble des indices correspondant aux variables apparaissant dans $\Phi(a)$. Pour tout t de $T(\Sigma,X_p)$ et tout p-uple $\vec{t}'=(t_1,\ldots,t_p)$, $t.\vec{t}'$ désigne l'arbre obtenu en substituant dans t, pour tout $i\leqslant p$, t_i à chaque occurrence de x_i. Nous disons que v est *facteur* de t si t se décompose en $u.v.\vec{w}$. Dans la suite, la notation $u.v.\vec{w}$ qui met en évidence un facteur v dans un arbre, suppose implicitement que $u\in\tilde{T}(\Sigma,X_1)$. L'ensemble $F(t)$ désigne tous les facteurs de t et si K est une forêt, $F(K)$ désigne l'union des $F(t)$ pour t dans K.

Nous supposons le lecteur familiarisé avec les automates d'arbres [10,8]. Un automate d'arbres est représenté par (Σ,Q,F,T) avec F l'ensemble des états de sortie, T l'ensemble des transitions.

3. Fidélité et morphismes

Dans cet article nous considérons des morphismes quasi-alphabétiques. Si Φ est un tel morphisme de $T(\Sigma,X)$ dans $T(\Delta,X)$, chaque image de lettre $f(x_1,\ldots,x_n)$ de Σ par Φ est de la forme x_i ou encore $u(x_{i_1},\ldots,x_{i_k})$ où $u\in CT_k(\Delta)$ et chaque variable apparaît à profondeur 1 dans u. En d'autres termes, Φ peut effacer f et le ≪remplacer≫ par un de ses fils, ou ≪transformer≫ f en insérant des sous-termes clos, en supprimant, copiant ou échangeant ses fils directs. Pour étudier les transformations par Φ, nous

[1] *Complet* est aussi appelé *nondeleting* par d'autres auteurs [9].

utiliserons des correspondances entre les facteurs d'un terme, et les facteurs de son image. Nous les énonçons dans la remarque suivante.

Remarque 1. Soit un v terme de $T(\Sigma)$ et soit p_1 une position dans $t=\Phi(v)$ et t_1 le facteur de t à cette position; alors si t_1 est de taille supérieure ou égale à $|\Phi|$, on peut trouver au moins une correspondance de p_1 dans v, i.e. il existe v_0, v_1, t_0 tels que $v_0.v_1=v$, $t_0.t_1=\Phi(v)=t$ et $\Phi(v_0)=t_0$, $\Phi(v_1)=t_1$, avec p_1 correspondant à une occurrence de variable dans t_0.

Fidélité. Rappelons qu'une relation τ de $U\times V$ est dite *fidèle* si pour tout t de V, $\tau^{-1}(t)$ est fini. Nous allons montrer que pour tout morphisme Φ et toute forêt reconnaissable R, on peut trouver R' reconnaissable inclus dans R tel que $\Phi(R)=\Phi(R')$ et (Φ,R') soit fidèle (on dit que (Φ,R') est fidèle si la relation $\{(t,\Phi(t))\,|\,t\in R'\}$ est fidèle).

Définition I. Φ est dit p-**fidèle** sur K si pour tout facteur u de K, $|u|$ est inférieure à $p\times(1+|\Phi(u)|)$.

Il se peut que (Φ,K) soit fidèle sans être p-fidèle pour aucun p (considérer $K=\{b(a^id,a^jd)\,|\,i\leqslant j\}$ et $\Phi(b(x,y))=y,\Phi(a(x))=a(x),\Phi(d)=d$). Par contre, si K est reconnaissable, alors les deux notions coïncident. Nous nous limitons ici à une simple proposition qui lie reconnaissabilité et p-fidélité.

Proposition 1. *Soit Σ,Δ deux alphabets gradués, Φ un morphisme de $T(\Sigma,X)$ dans $T(\Delta,X)$ et K une forêt reconnaissable sur l'alphabet Σ. On peut construire une forêt reconnaissable K' incluse dans K tel que $\Phi(K)=\Phi(K')$ et telle que (Φ,K') soit p-fidèle.*

Démonstration. Nous allons construire une forêt K' à partir de K en ôtant les termes de K suffisamment grands, dont l'image peut être obtenue par application de Φ à des termes plus petits. On distingue deux ensembles qui vont traiter les cas des règles effaçantes de la forme $\Phi(a)=x_i$, et ceux des règles non complètes de la forme $\Phi(f)=u(x_{i_1},\dots,x_{i_n})$ pour f d'arité plus grande que n.

Soit

$$E_i = \{t\,|\,\forall u\in F(t)\ |u|>i\Rightarrow|\Phi(u)|>0\}$$

$$C_i = \{t\,|\,\forall u\in F(t),\ \forall f\in\Sigma,$$
$$[t=u.f(u_1,\dots,u_n)\wedge x_j\notin Var(\Phi(f))]\Rightarrow|u_j|\leqslant i\}$$

E_i est l'ensemble des termes qui ne contiennent pas de facteurs effaçables par Φ de taille supérieure à i. L'ensemble C_i désigne les termes pour lesquels seuls des sous-termes de hauteur inférieure à i peuvent être élagués par des règles non-complètes.

Soit $A=(\Sigma,Q,F,T)$ un automate déterministe ascendant qui reconnaît K. Soit $p=2\times|Q|$ ($|Q|$ est le cardinal de Q). Soit $K'=K\cap E_p\cap C_p$.

Nous montrons maintenant que $\Phi(K')=\Phi(K)$. Puisque $K'\subseteq K$, il suffit de montrer $\Phi(K')\supseteq\Phi(K)$. Soit t dans $\Phi(K)$, soit v le plus petit terme (en nombre de noeuds) de K tel que $\Phi(v)=t$. Nous montrons que v appartient à la fois à E_p et C_p et donc à K'.

- Si v n'est pas dans C_p alors $v=u.f(u_1,\dots,u_n)$ et un sous-terme u_i de hauteur supérieure à $p=2\times|Q|$ est effacé[2] par Φ. Soit q l'état associé par A à u_i. Sur une plus longue branche de u_i, on trouve nécessairement deux occurrences d'un même état de A. On peut alors appliquer le lemme de l'étoile dans les arbres pour construire un terme u_i' plus petit que u_i et dont l'état associé par A est q. On obtient alors une contradiction puisque $u.f(u_1,\dots,u_i',\dots,u_n)$ est dans K, plus petit que v et de même image t par Φ.
- Si v n'est pas dans E_p, c'est qu'il existe un facteur u de v, de taille supérieure à p et effacé par Φ. Alors v peut s'écrire sous la forme $v=v'.u.v''$ tel que $t=\Phi(v)=\Phi(v')$. $\Phi(v'')$. Considérons un chemin π dans u de longueur maximale et la position ρ de u à laquelle v'' apparaît dans v. Alors $\pi=\pi_1.\pi_2$ où π_1 est l'ensemble des positions inférieures à la position ρ. Puisque la longueur de π est supérieure à $2\times|Q|$, soit π_1 soit π_2 est de longueur supérieure à $|Q|$. Comme dans le premier cas, nous pouvons trouver en appliquant le lemme de l'étoile dans les arbres un terme u_i' plus petit que u_i pour lequel A associe le même état qu'à u_i et tel que $v'.u'.v''$ soit d'image t par Φ. D'où la contradiction.

Enfin, nous montrons par récurrence sur la taille de tout facteur u de K' que Φ est bien p-fidèle sur K', donc que $|u|\leqslant p(1+|\Phi(u)|)$.

- Soit $u\in K'$, $|u|<p$. Il est facile de voir que même si $|\Phi(u)|=0$, alors $|u|\leqslant p(1+|\Phi(u)|)$.
- Soit $u\in K'$, $|u|\leqslant 2\times p$. Si $|u|>p$ alors $|\Phi(u)|>1$ puisque u est dans E_p. Donc, $|u|\leqslant p(1+|\Phi(u)|)$.
- Supposons que Φ est bien p-fidèle pour l'ensemble des facteurs de K' de taille inférieure à k et soit un facteur u de K' de taille $k+1$. Remarquons que u peut s'écrire $w.w'$ avec $|u|=p+|w'|$. Puisque $t\in E_p$, nous avons

$$|\Phi(w)|\geqslant 1. \tag{1}$$

De plus, puisque $|w'|>p$ et que w' est un facteur de t dans C_p, alors w' n'est pas élagué par Φ. On obtient donc

$$\begin{aligned} p(1+|\Phi(u)|) &\geqslant p(1+1+|\Phi(w')|) && \text{(d'après (1))}\\ &\geqslant p+p(1+|\Phi(w')|) \\ &\geqslant p+|w'| && \text{(par hypothèse d'induction}\\ &\geqslant |u|. \quad \square \end{aligned}$$

[2] Pour la condition d'appartenance à C_p, on pourrait se limiter à $p=|Q|$.

4. Non-linéarité et reconnaissabilité

Soit Φ un morphisme et K un reconnaissable. Si Φ est linéaire, $\Phi(K)$ est reconnaissable; la réciproque est bien sûr fausse, comme l'illustre l'exemple suivant. Soit le morphisme non linéaire

$$\begin{aligned} \Phi: \quad a(x) &\rightarrow c(x,x) \\ b(x,y) &\rightarrow b(x,y) \\ d &\rightarrow d \end{aligned}$$

et K l'ensemble de tous les arbres clos construits à partir de b, d, et $a(d)$). Alors $\Phi(K)$ est l'ensemble (reconnaissable) de tous les termes construits à partir de b, d, et $c(d,d)$.

Cependant la proposition suivante correspond à une sorte de réciproque faible. Certes, il n'existe pas d'ensemble K' reconnaissable et inclus dans K tel que $\Phi(K)=\Phi(K')$ et tel que Φ soit linéaire sur K'. Néanmoins, nous pouvons dire informellement que presque tout peut être obtenu de façon linéaire. Ainsi, dans l'exemple ci-dessus, le seul sous-arbre qui ne peut être obtenu de façon linéaire est $c(d,d)$ et si nous définissons $R=T_{\{a',b,d\}}$ et

$$\begin{aligned} \pi: b(x,y) &\rightarrow b(x,y) \\ d &\rightarrow d \\ a' &\rightarrow c(d,d) \end{aligned}$$

alors $\Phi\circ\pi$ est linéaire, $\pi(R)\subset K$ et $\Phi\circ\pi(R)=\Phi(K)$. La proposition suivante traduit cette idée. Elle suppose le morphisme Φ quasi-alphabétique mais nous conjecturons qu'elle reste vraie dans le cas non-alphabétique.

Proposition 2. *Soit Φ un morphisme quasi-alphabétique de $T(\Sigma,X)$ dans $T(\Delta,X)$ et K une forêt reconnaissable sur Σ. Si $\Phi(K)$ est reconnaissable alors il existe une forêt reconnaissable R et un morphisme π linéaire, quasi-alphabétique complet et strict tel que $\Phi\circ\pi$ soit linéaire, $\pi(R)\subset K$ et $\Phi\circ\pi(R)=\Phi(K)$.*

La preuve de cette proposition se décompose ainsi : dans un premier temps nous exprimons que tout terme dans $\Phi(K)$, sauf un nombre fini, peut être obtenu sans pour autant que la non-linéarité du morphisme soit nécessaire. Pour cela, nous définissons la condition $\mathbf{C}_n(t_0,\vec{t})$ qui exprime que le terme $t_0.\vec{t}$ de $\Phi(K)$ peut être obtenu en construisant chaque t_i de $\vec{t}$ indépendamment (voir Fig. 1). L'ensemble $\heartsuit$ des termes dans $\vec{t}$ (pour tout t_0) vérifiant la condition $\mathbf{C}_n(t_0,\vec{t})$ est ensuite montré cofini (Lemme 3). Pour prouver ceci, nous construisons cet ensemble en utilisant le fait que K est reconnaissable. La preuve repose sur une induction sur le nombre n de variables dans le contexte t_0.

Dans un deuxième temps, nous utilisons la propriété de p-fidélité du morphisme Φ et la cofinitude de $\heartsuit$ pour définir une extension (finie) de l'alphabet Σ qui encode en une simple quasi-lettre, tout terme de hauteur trop petite pour ne pas avoir l'assurance que son image est dans $\heartsuit$. Cette transformation légitime l'apparition du morphisme π.

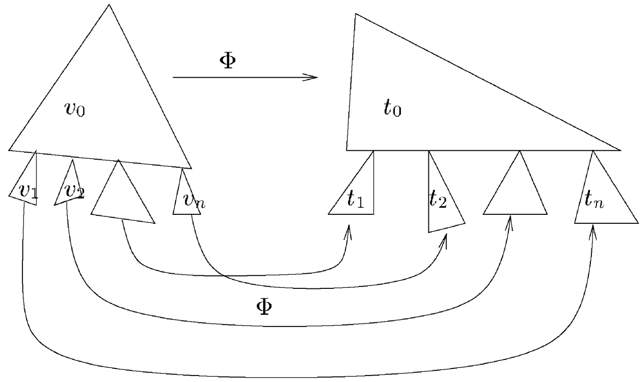

Fig. 1. La condition $\mathbf{C}_n(t_0, t)$. Construire d'abord les termes de t de façon indépendante.

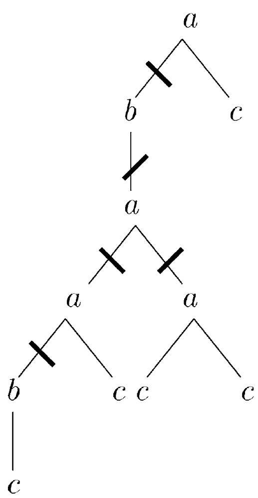

Fig. 2. Le découpage selon dec de $a(b(a(a(b(c),c),a(c,c))),c))$. Les éléments dans Ter sont $a(c,c)$ et $b(c)$.

La phase finale de la preuve définit un découpage en couches de tout terme image dans $\Phi(K)$ de la racine vers les feuilles (Fig. 2). La présence de termes de $\heartsuit$ dans la couche la plus basse (vers les feuilles) de ce découpage, conjuguée avec la condition $\mathbf{C}_n(.,.)$ va permettre d'exhiber une construction ≪linéaire≫ de cette image.

5. Preuve de la Proposition 2

Montrons d'abord que nous pouvons supposer que $\Phi(K) = T(\Delta)$. En effet, supposons que la proposition soit vraie avec cette restriction. Maintenant, soit (Φ, K) tel que $\Phi(K)$ soit reconnaissable. Soit $\bar{\Delta}$ une copie de Δ et $i : \bar{\Delta} \to \Delta$ l'injection naturelle. On peut définir $(\bar{\Phi}, \bar{K})$ par:

- $\bar{K} = K \cup i^{-1}({}^c\Phi(K))$ où ${}^cA$ désigne le complémentaire de A dans $T(\Delta)$
- $\bar{\Phi} : \bar{\Phi}_{/\Sigma} = \Phi, \bar{\Phi}_{/\bar{\Delta}} = i$;

$\bar{K}$ est reconnaissable; $\bar{\Phi}$ conserve les propriétés (être fidèle et quasi-alphabétique) de Φ et $\bar{\Phi}(\bar{K}) = T(\Delta)$. Donc, d'après notre hypothèse, la proposition s'applique à $\bar{\Phi}$ et $\bar{K}$ et nous pouvons construire π et $\bar{R}$ qui satisfont les conditions de la proposition. Considérons maintenant $R = \bar{R} \cap \pi^{-1}(K)$; le lecteur peut alors vérifie que la proposition est vraie aussi pour (Φ, π, R).

Nous supposons donc, à partir de maintenant, que Φ est quasi-alphabétique, (Φ, K) fidèle et $\Phi(K) = T(\Delta)$.

La condition $\mathbf{C}_n(t_0, \vec{t})$ est essentielle dans notre construction. Dans le Lemme 3 dit du coeur, notre raisonnement par induction repose sur la possibilité d'$\ll$étendre$\gg$ cette condition de n à $n+1$. L'expression de la condition $\mathbf{CF}(t_0, t)$ et ses propriétés (Lemme 1) nous permet de le faire.

Condition $\mathbf{CF}(t_0, t)$. Soient $t_0 \in CT_1(\Delta)$ et $t \in T(\Delta)$. La condition $\mathbf{CF}(t_0, t)$ est définie par:

$$\forall v \in K, \quad \left(\Phi(v) = t_0.t \Rightarrow \begin{array}{l} \exists v_0 \in CT_1(\Sigma) \\ \exists v_1 \in T(\Sigma) \end{array} \Big/ \begin{array}{rcl} \Phi(v_0) & = & t_0 \\ \Phi(v_1) & = & t \\ v & = & v_0.v_1 \end{array} \right). \qquad (\mathbf{CF}(t_0, t))$$

Intuitivement, $\mathbf{CF}(t_0, t)$ signifie que même si $t_0.t$ peut être obtenu d'une façon non linéaire, t est toujours obtenu indépendamment de tout sous-terme de t_0.

Condition $\mathbf{C}_n(t_0, \vec{t})$. Soit n un entier et $t_0 \in CT_n(\Delta)$. Soit $\vec{t}$ un n-uple de termes dans $T(\Delta)^n$. La condition $\mathbf{C}_n(t_0, \vec{t})$ est définie par:

$$\exists v_0 \in \widetilde{CT}_n(\Sigma), \exists \vec{v} \in T(\Sigma)^n \Big/ \begin{array}{rcl} v_0.\vec{v} & \in & K \\ \Phi(v_0) & = & t_0 \\ \Phi(\vec{v}) & = & \vec{t}. \end{array} \mathbf{C}_n(t_0, t) \qquad (\mathbf{C}_n(t_0, t))$$

La condition $\mathbf{C}_n(t_0, \vec{t})$ est moins restrictive. Intuitivement, cela signifie qu'il y a au moins une façon d'obtenir $\vec{t}$ de façon linéaire dans $t_0.\vec{t}$, i.e. telle que chaque composante de $\vec{t}$ soit obtenue indépendamment d'une part des autres composantes, d'autre part des sous-termes de t_0.

Comme $\Phi(K) = T(\Delta)$, $\mathbf{CF}(t_0, t)$ implique de façon évidente $\mathbf{C}_n(t_0, t)$ pour tout couple (t_0, t) de $CT_1(\Delta) \times T(\Delta)$. Le lemme suivant donne une condition suffisante pour $\mathbf{CF}(t_0, t)$:

Lemme 1. *Si* $|t| > \max(|\Phi|, |t_0|)$, $\mathbf{CF}(t_0, t)$ *est vérifiée.*

Démonstration. Si on pense à la signification de $\mathbf{CF}(t_0, t)$, le lemme est évident : si la profondeur de t est supérieure à $|\Phi|$, t ne peut-être un sous-terme de l'image d'une lettre ; si de plus $|t|$ est supérieure à $|t_0|$, t ne peut être égal à un sous-terme de t_0, et donc est nécessairement obtenu indépendamment de tout sous-terme de t_0.

Formellement, soit v tel que $\Phi(v)=t_0.t$. Par la remarque 1, comme Φ est quasi-alphabétique et que la profondeur de t est supérieure à $|\Phi|$, on peut trouver (v_0, v_1) dans $CT_1(\Delta)_1 \times T(\Delta)$ tel que:

- $v_0.v_1 \in K$;
- $\Phi(v_0.v_1)=t_0.t$;
- $\Phi(v_0)=u.(x,\dots,x)$, $u \in CT_p(\Delta)$, $p \geqslant 1$;
- $\Phi(v_1)=t$.

Alors $t_0.t=u(t,\dots,t)$ pour $u \in CT_p(\Delta), p \geqslant 1$. Si $p>1$ alors t est un sous-terme de t_0. Or, $|t|>|t_0|$ donc $p=1$. La condition $\mathbf{CF}(t_0,t)$ est donc vérifiée. □

Nous utilisons cette propriété pour prouver le lemme suivant:

Lemme 2. *Soit un n-uple de termes* $\vec{t}=(t_1,\dots,t_n)$:

$$\left.\begin{array}{r}\mathbf{C}_n(t_0.(x_1,\dots,x_n,t_{n+1}),\vec{t})\\ \mathbf{CF}(t_0.(\vec{t},x_{n+1}),t_{n+1})\end{array}\right\} \Rightarrow \mathbf{C}_{n+1}(t_0,(\vec{t},t_{n+1})).$$

Démonstration. Remarquons d'abord que $t_0.(\vec{t},t_{n+1})=t_0.(x_1,\dots,x_n,t_{n+1}).\vec{t}$. Donc, si la condition $\mathbf{C}_n(t_0.(x_1,\dots,x_n,t_{n+1}),\vec{t})$ est vérifiée, il existe au moins une façon d'obtenir $t_0.(\vec{t},t_{n+1})$ par Φ telle que chaque composante de $\vec{t}$ est obtenue indépendamment des autres et de chaque sous-terme de t_0, i.e.:

$$\exists v_0, \exists \vec{v} \text{ tels que } \begin{cases} \Phi(v_0) &= t_0.(x_1,\dots,x_n,t_{n+1}),\\ \Phi(\vec{v}) &= \vec{t},\\ v_0.\vec{v} &\in K.\end{cases}$$

Mais on a aussi $t_0(\vec{t},t_{n+1})=t_0.(\vec{t},x_{n+1}).t_{n+1}$ et donc par $\mathbf{CF}(t_0.(\vec{t},x_{n+1}),t_{n+1})$, t_{n+1} est obtenu indépendamment de chaque sous-terme de $t_0.(\vec{t},x_{n+1})$, i.e.:

$$\exists w_0, w_1 \text{ tels que } \begin{cases} \Phi(w_0) &= t_0.\vec{t},\\ \Phi(w_1) &= t_{n+1},\\ w_0.w_1 &= v_0.\vec{v}.\end{cases}$$

Mais alors, $\Phi(w_0 \wedge v_0)=t_0$ (avec $w_0 \wedge v_0 = \inf\{w_0, v_0\}$ pour l'ordre $u<v$ si $v=u.w$) et $\mathbf{C}_{n+1}(t_0,(\vec{t},t_{n+1}))$ est vérifiée. □

Nous énonçons maintenant le lemme central (le lemme du coeur) qui exprime que les branches qui ne peuvent pas être obtenues de façon linéaire appartiennent à une forêt finie. L'idée de la preuve est la suivante: d'après ce qui précède, pour tout arbre t de $CT_1(\Delta)$, $\mathbf{CF}(t,v)$ est vérifiée pour presque tout v. En utilisant le fait que $\Phi(K)=T(\Delta)$, nous montrons qu'il existe une forêt cofinie, i.e. dont le complémentaire dans $T(\Delta)$ est fini, telle que pour tout arbre t de $CT_1(\Delta)$, pour tout v de cette forêt, $\mathbf{CF}(t,v)$ est vérifiée. Par récurrence, nous prouvons ensuite:

Lemme 3 (du coeur). *On peut construire une forêt cofinie sur Δ qu'on notera $\heartsuit$, telle que pour tout entier n, pour tout arbre t_0 de $CT_n(\Delta)$, et pour tout vecteur $\vec{t}$ de $\heartsuit^n$, $\mathbf{C}_n(t_0,\vec{t})$ est vérifiée.*

Démonstration. Soit $A=(\Sigma,Q,F,T)$ un automate ascendant qui reconnaît K. Soit $S_{Q'}$ la forêt reconnue par l'automate ascendant (Σ,Q,Q',T). Notre forêt $\heartsuit$ sera l'intersection de toutes les forêts $\Phi(S_R)$ pour toute partie R de Q telle que $\Phi(S_R)$ soit cofinie. Comme Q est fini et que chaque forêt $\Phi(S_R)$ est cofinie, $\heartsuit$ est bien cofinie.

- $n=1$.
  Soit t_0 de $CT_1(\Delta)$. Soit Q' l'ensemble des états q tels qu'il existe un arbre v_0 de $CT_1(\Sigma)$ avec $v_0(q)\xrightarrow{*}f$, avec $f\in F$ et $\Phi(v_0(x))=t_0$. Soit la forêt $\mathscr{C}_{t_0}$ définie par $\mathscr{C}_{t_0}=_{def}\{t\in T(\Delta)\mid \mathbf{CF}(t_0,t)\}$. D'après le Lemme 1, $\mathscr{C}_{t_0}$ contient $\{t\in T(\Delta)\mid |t|>\max(|t_0|,|\Phi(\alpha)|)\}$ et est donc cofinie. Par construction, $\Phi(S_{Q'})=\mathscr{C}_{t_0}$ et donc est cofinie. Par conséquent, pour tout t_0, $\mathscr{C}_{t_0}$ contient $\heartsuit$. Donc, pour tout t_0 de $CT_1(\Delta)$, chaque arbre de $\heartsuit$ appartient à $\mathscr{C}_{t_0}$, i.e. pour chaque arbre t_0 de $CT_1(\Delta)$, pour chaque t de $\heartsuit$, $\mathbf{CF}(t_0,t)$ est vérifiée, donc $\mathbf{C}_1(t_0,t)$ l'est aussi.
- pas de récurrence: supposons la propriété vraie pour n.
  Soit t_0 dans $CT_{n+1}(\Delta)$, $\vec{t}$ dans $\heartsuit^n$. Par hypothèse de récurrence, pour chaque arbre t_{n+1} de $T(\Delta)$, $\mathbf{C}_n(t_0.(x_1,\ldots,x_n,t_{n+1}),\vec{t})$ est vérifiée. Donc, par les Lemmes 1 et 2, si la profondeur de t_{n+1} est supérieure à celle de $t_0.(\vec{t},x_{n+1})$ et à la profondeur maximale de l'image d'une lettre, $\mathbf{C}_{n+1}(t_0,(\vec{t},t_{n+1}))$ est assurée. Soit maintenant Q'' la partie de Q définie par:

$$Q''=\left\{ q\in Q \;\middle|\; \begin{array}{l} \exists i,\ \exists v_0\in CT_{n+1}(\Sigma),\exists(q_1,\ldots,q_n)\in Q^n \text{ tels que:}\\ v_0(q_1,\ldots,q_i,q,q_{i+1},\ldots,q_n)\xrightarrow{*}F\\ \Phi(v_0(x_1,\ldots,x_i,x,x_{i+1},.,x_n))=t_0.(x_{i_1},\ldots,x_{i_n},x)\\ \forall j\neq k, i_j\neq i_k\\ \forall j, t_j\in\Phi(S_{q_{i_j}}) \end{array} \right\}.$$

  Alors, t_{n+1} appartient à $\Phi(S_{Q''})$ si et seulement si $\mathbf{C}_{n+1}(t_0,(\vec{t},t_{n+1}))$ est vérifiée; donc $\Phi(S_{Q''})$ est cofinie et contient $\heartsuit$ par construction; d'où $\mathbf{C}_{n+1}(t_0,(\vec{t},t_{n+1}))$ est vérifiée pour tout t_{n+1} de $\heartsuit$. □

Par ce lemme on est maintenant en mesure de construire π et R satisfaisant aux conditions de la Proposition 2.

Soit $k=\max_{t\notin\heartsuit}|t|+1$ (Comme $\heartsuit$ est cofini, k est bien défini). Comme (Φ,K) est fidèle, nous savons d'après la première partie qu'il existe un entier p tel que (Φ,K) soit p-fidèle. Soit la forêt P définie par:

$$P=_{def}\left\{ s\in CT_1(\Sigma) \;\middle|\; \begin{array}{l} |s|\leqslant p\times(1+k)\\ \Phi(s) \text{ est une quasi-lettre}\\ \Phi(s) \text{ est linéaire} \end{array} \right\}.$$

Soit maintenant un nouvel alphabet Σ_P en bijection avec $P : \Sigma_P =_{def} \{\alpha_s \mid s \in P\}$, α_s ayant l'arité de s. Alors π désigne le morphisme défini par $\pi(\alpha_s) = s$, pour tout α_s dans P et on définit R par $R = \pi^{-1}(K)$. Alors:

a. Par construction, π est linéaire, complet, quasi-alphabétique, strict et $\pi(R)$ est inclus dans K. De plus, $\Phi \circ \pi$ est linéaire par définition même de P.
b. II faut montrer que $\Phi \circ \pi(R) = T(\Sigma)$; ici réside la difficulté de la preuve. Intuitivement il s'agit de montrer que les sous-arbres qu'on ne peut pas obtenir par Φ de façon linéaire et qui sont donc, d'après le Lemme 3, le complément de $\heartsuit$, sont images d'arbres de P. Nous allons pour cela définir pour chaque arbre t de $T(\Delta)$ un découpage noté $dec(t)$ tel que chaque branche ≪terminale≫ de t soit de profondeur k et donc appartienne à $\heartsuit$.

Nous définissons dec par:

$$
\begin{aligned}
dec(a.(t_1,\ldots,t_n)) &= (\alpha, dec(u_1),\ldots,dec(u_n)) \text{ avec :}\\
&\quad a.(t_1,\ldots t_n) = \alpha.(u_1,\ldots u_n)\\
&\quad \alpha = a.(v_1,\ldots v_n) \text{ avec } v_i = t_i \text{ si } |t_i| < k,\ v_i = t_i \text{ sinon.}
\end{aligned}
$$

Par exemple pour $k=2$ avec $\Sigma_0 = \{c\}$, $\Sigma_1 = \{b\}$, $\Sigma_2 = \{a\}$, on obtient:

$$
\begin{aligned}
&dec(a(b(a(a(b(c),c),a(c,c))),c))\\
&\quad = (a(x_1,c),(b(x_1),(a(x_1,x_2),(a(x_1,c),b(c)),a(c,c)))).
\end{aligned}
$$

On peut remarquer que, pour tout terme t, chaque élément de $dec(t)$ est soit un terme clos de hauteur k, soit une quasi-lettre de hauteur au plus k.

Maintenant, soit t un terme de $T(\Delta)$. On va montrer qu'il existe un terme v de R tel que $\Phi \circ \pi(v) = t$. On distinguera deux cas:

- $|t| \leqslant k$. Soit alors u dans K tel que $\Phi(u) = t$; comme Φ est p-fidèle, u appartient à P et donc α_u appartient à R avec $\Phi \circ \pi(\alpha_u) = t$.
- $|t| > k$. Alors chaque élément de $dec(t)$ est de profondeur inférieure ou égale à k. De plus, on peut décomposer t en $t = to.\vec{t}$ avec soit $\vec{t}$ correspondant aux éléments clos du découpage: chaque composante de $\vec{t}$ est de profondeur k et donc appartient à $\heartsuit$. Donc, par le Lemme 3, $\mathbf{C}_n(t_0,\vec{t})$ est vérifiée: on peut trouver v_0 dans $\widehat{CT}_n(\Sigma)$, et $\vec{v}$ dans $T(\Sigma)^n$ tels que $v_0.\vec{v}$ appartienne à K et $\Phi(v_0) = t_0, \Phi(\vec{v}) = \vec{t}$.
  On va montrer qu'on peut trouver un découpage de $v_0.\vec{v}$ qui corresponde exactement par Φ au découpage $dec(t)$ et tel que chaque élément du découpage de v appartienne à P.
  Remarquons d'abord que par construction les composantes de $\vec{v}$ sont dans P puisqu'elles ont pour image des terme clos de hauteur k et que (Φ,K) est p-fidèle.
  Il suffit donc maintenant de trouver une décomposition de v_0 correspondant par Φ à $dec(t)$. Raisonnons en fonction de $n(t)$, le nombre d'éléments non clos $dec(t)$. Si $n(t) = 1$, le découpage sera simplement v_0. Comme $\mathbf{C}_n(t_0,\vec{t})$ est vérifiée, $\Phi(v_0) = t_0$ et est donc une quasi-lettre linéaire. De plus, comme (Φ,K) est p-fidèle, $|v_0| \leqslant p\times(1+k)$.

Si $n(t)>1$, on peut d'abord remarquer que puisque Φ est quasi-alphabétique et que les éléments du découpage sont des quasi-lettres ou des termes clos de hauteur supérieure à $|\Phi|$, on peut décomposer v_0 en $\beta_0(w_1,\ldots,w_k)$, $\Phi(\beta_0)$ correspondant au premier élément de $dec(t)$ aux variables près. De plus, comme $\mathbf{C}_n(t_0,\vec{t})$ est vérifiée, $\Phi(\beta_0)$ est linéaire car chacune de ses variables est remplacée dans t par un terme qui contient une composante de $\vec{t}$. En itérant le processus sur les termes w_i, on peut définir un découpage de v_0 tel que chaque élément du découpage ait pour image un élément du découpage de t (donc une quasi-lettre linéaire). De plus, comme (Φ,K) est p-fidèle, et que chaque élément de $dec(t)$ est un élément de $F(K)$, chaque élément du découpage a une hauteur au plus $p\times(1+k)$ et donc est dans P.

Donc $v_0.\vec{v}$ appartient à appartient à $\pi\circ\pi^{-1}(K)$ et t appartient à $\Phi\circ\pi(R)$.

Remarque. Même si Φ est alphabétique, π peut ne pas être alphabétique; par exemple, soit Φ : $a(x)\to b(x,x), c\to c$, $K=a(c)$. Supposons qu'il existe R reconnaissable et Π linéaire, *alphabétique*, complet et strict tel que $\Phi\circ\Pi$ soit linéaire, $\Pi(R)\subset K$ et $\Phi\circ\Pi(R)=\Phi(K)$; alors, soit t tel que $\Phi\circ\Pi(t)=b(c,c)$; nécessairement, comme $\Phi\circ\Pi$ est complet, alphabétique, linéaire et strict, la taille de t est 3; mais alors, la taille de $\Pi(t)$ est 3 également et $\Pi(t)$ n'appartient pas à K.

Nous conjecturons que cette proposition est vraie même si Φ n'est pas quasi-alphabétique (et alors bien sûr $\Phi\circ\Pi$ ne le serait pas non plus). La preuve donnée ici ne s'étend malheureusement pas, le ≪découpage≫ de l'image d'un arbre étant beaucoup plus complexe dans le cas général.

Toutes les constructions ci-dessus sont effectives; en effet, pour un morphisme quasi-alphabétique, on peut décider si l'image d'un reconnaissable est reconnaissable et si elle est cofinie (ce résultat est obtenu en utilisant une classe d'automates à tests [3]). On peut encore souligner le saut dans la difficulté quand on essaie de généraliser aux morphismes non-quasi-alphabétiques: Le problème de la décidabilité de la reconnaissabilité de l'image d'un reconnaissable par un morphisme quelconque est ouvert.

Cette proposition nous permet d'étendre le **théorème de transversale rationnelle** (ou cross-section) dans les arbres. Rappelons le théorème de cross-section: si Φ est un morphisme linéaire de $T(\Sigma)$ dans $T(\Delta)$, toute forêt reconnaissable F contient une partie reconnaissable qui est en bijection par Φ avec $\Phi(F)$. A. Arnold a montré que l'hypothèse Φ linéaire ne pouvait être supprimée [1]. Nous conjecturons qu'elle peut être remplacée par ≪$\Phi(F)$ reconnaissable≫, i.e.:

Conjecture 1. *Soit Φ un morphisme, F une forêt reconnaissable telle que $\Phi(F)$ soit reconnaissable. alors F contient une forêt reconnaissable qui est en bijection par Φ avec $\Phi(F)$.*

Nous montrons ici cette conjecture dans le cas où Φ est quasi-alphabétique:

Théorème 1 (Théorème de transversale rationnelle). *Soit Φ un morphisme quasi-alphabétique, R une forêt reconnaissable tels que $\Phi(R)$ est reconnaissable. Alors, il existe une forêt reconnaissable R' incluse dans R tel que Φ soit une bijection de R' sur $\Phi(R)$.*

Démonstration. La preuve est simple et utilise le théorème de cross-section pour les morphismes linéaires; en effet, soit Φ un morphisme quasi-alphabétique et R une forêt reconnaissable telle que $\Phi(R)$ soit reconnaissable. On peut donc trouver π et K satisfaisant aux conditions de la Proposition 2. Comme $\Phi\circ\pi$ est linéaire, on peut appliquer le théorème de cross-section à $\Phi\circ\pi$ et K. Donc, il existe K' reconnaissable inclus dans K tel que $\Phi\circ\pi(K')=\Phi\circ\pi(K)=\Phi(R)$ et tel que $\Phi\circ\pi$ soit injectif sur K'. Soit $R'=\pi(K')$; comme $\pi(K)$ est inclus dans R, R' est bien inclus dans R; comme π est linéaire, R' est reconnaissable; comme $\Phi\circ\pi$ est injectif sur K', Φ est injectif sur R'. $\square$

6. Bimorphismes

Dans ce paragraphe, nous appliquons les résultats précédents aux bimorphismes d'arbres. Un bimorphisme est un triplet $\tau=(\Phi,K,\Psi)$, Φ,Ψ étant des morphismes et K un reconnaissable, et la transformation associée est définie par $\tau(t)=\Psi(\Phi^{-1}(t)\cap K)$. Dans le cas des mots, les transformations définies par bimorphismes sont exactement les transductions rationnelles [16], dans le cas des arbres la situation est plus complexe [2].

Dans ce paragraphe, nous montrons que sous certaines conditions, on peut supposer que les composantes d'un bimorphisme ont certaines propriétés (linéarité, complétude, alphabécité,...). En particulier, le paragraphe précédent nous permet de réduire la non-linéarité sous des hypothèses de reconnaissabilité:

La Proposition 2 nous permet d'obtenir les résultats suivants concernant les bimorphismes:

Lemme 4. *Soit $\tau=(\Phi,K,\Psi)$ un bimorphisme* **fonctionnel** *avec Φ quasi-alphabétique. Si le domaine de τ est reconnaissable, on peut construire $\tau'=(\Phi',K',\Psi')$ équivalent à τ avec*:
- Φ' **linéaire**;
- Φ' *quasi-alphabétique*; (*mais l'éventuelle alphabécité de Φ n'est pas conservée*).
- Φ' (*resp.* Ψ') *strict* (*resp. complet*) *si Φ* (*resp.* Ψ) *l'était*;
- Ψ' *quasi-alphabétique si Ψ l'était*; (*mais l'éventuelle alphabécité de Ψ n'est pas conservée*).

Démonstration. La preuve est immédiate; d'après la Proposition 2, on peut construire un reconnaissable R et un morphisme π linéaire, quasi-alphabétique complet et strict tel que $\Phi\circ\pi$ soit linéaire, $\pi(R)\subset K$ et $\Phi\circ\pi(R)=\Phi(K)$. Alors soit $\tau'=(\Phi'=\Phi\circ\pi, K'=R, \Psi'=\Psi\circ\pi)$: vu les propriétés de π et de $\Phi\circ\pi$, Π' et Ψ' vérifient les conditions requises; soit t dans le domaine de τ; comme $\pi(K')\subset K$, il est immédiat que $\tau'\subset\tau$ (au sens de l'inclusion des relations); comme $\Phi\circ\pi(K')=\Phi(K)$, τ et τ' ont même domaine; comme de plus τ est fonctionnel, on a bien τ' équivalent à τ.

Symétriquement, on a:

Lemme 5. *Soit $\tau=(\Phi,K,\Psi)$ un bimorphisme* $\ll$injectif$\gg$—i.e.. $\tau(t)\cap\tau(t')=\emptyset$, *si $t\neq t'$ — avec Ψ quasi-alphabétique. Si l'image de τ est reconnaissable, on peut construire*

$\tau' = (\Phi', K', \Psi')$ équivalent à τ avec:
- Ψ' **linéaire**;
- Ψ' *quasi-alphabétique*; (*mais l'éventuelle alphabécité de Ψ n'est pas conservée*).
- Ψ' (*resp.* Φ') *strict* (*resp. complet*) *si* Ψ (*resp.* Φ) *l'était*;
- Φ' *quasi-alphabétique si Φ l'était*; (*mais l'éventuelle alphabécité de Φ n'est pas conservée*).

Démonstration. Il suffit d'appliquer la Proposition 2 à (Ψ, K).

La conjonction des deux lemmes précédents donne:

Théorèm 2. *Soit une bijection de domaine et d'image reconnaissables réalisée par un bimorphisme $\tau = (\Phi, K, \Psi)$ avec Φ et Ψ quasi-alphabétiques. On peut alors réaliser cette bijection par un bimorphisme $\tau' = (\Phi', K', \Psi')$ avec*:
- Φ' et Ψ' **linéaires**;
- Φ' *et* Ψ' *quasi-alphabétiques*; (*mais l'éventuelle alphabécité de Φ ou Ψ n'est pas conservée*).
- Φ' (*resp.* Ψ') *strict* (*resp. complet*) *si* Φ (*resp.* Ψ) *l'était*.

7. Conclusion

L'approche développée ici face aux problèmes de non-linéarité est complémentaire de celle développée par exemple dans [4,5] où sont définies de nouvelles classes d'automates pouvant reconnaître certaines images homomorphes non-linéaires de reconnaissables. L'étude des homomorphismes ni linéaires ni quasi-alphabétiques semble plus difficile; certaines propriétés obtenues ici dans le cas quasi-alphabétique sont sans doute généralisables mais les techniques de preuve utilisées ici ne semblent pas s'étendre. Pour finir signalons un autre problème résolu dans le cas quasi-alphabétique [3] mais ouvert dans le cas général: *Peut-on décider si l'image homomorphe d'une forêt reconnaissable est reconnaissable?*

Rèfèrences

[1] A. Arnold, Le théorème de transversale rationnelle dans les arbres, Tech. Report I.T.3-78, Laboratoire d'Informatique Fondamentale de Lille, 1978.

[2] A. Arnold, M. Dauchet, Morphismes et bimorphismes d'arbres, Theoret. Comput. Sci. 20 (1982) 33–93.

[3] B. Bogaert, F. Seynhaeve, S. Tison, The recognizability problem for tree automata with comparisons between brothers, in: Proc. ETAPS 99 (FOSSACS), Lecture Notes in Computer Science, Vol. 1578, Springer, Berlin, 1999, pp. 150–164.

[4] B. Bogaert, S. Tison, Equality and disequality constraints on direct subterms in tree automata, in: Proc. 9th Annual Symposium on Theoretical Aspects of Computer Science, Vol. 577, Springer, Berlin, 1992, pp. 161–172.

[5] A.-C. Caron, H. Comon, J.-L. Coquidé, M. Dauchet, F. Jacquemard, Pumping, cleaning and symbolic constraints solving, in: S. Abiteboul, E. Shamir (Eds.), Automata, Languages and Programming, 21st Internat. Colloq., Lecture Notes in Computer Science, Vol. 820, Springer, Berlin, Jerusalem, Israel, 11–14 July, 1994, pp. 436–449.

[6] A.-C. Caron, J.-L. Coquidé, M. Dauchet, Encompassment properties and automata with constraints, in: Proc. 5th Internat. Conf. on Rewriting Techniques and Applications, Vol. 690, 1993, pp. 328–342.

[7] H. Comon, Equational formulas on order-sorted algebras, in: Proc. Internat. Colloq. on Automata Languages and Programming, Lecture Notes in Computer Science, Vol. 443, Springer, Berlin, 1990, pp. 674–688.

[8] H. Comon, M. Dauchet, R. Gilleron, D. Lugiez, S. Tison, M. Tommasi, Tree automata techniques and applications, 1997, available on: http://www.grappa.univ-lille3.fr/tata

[9] J. Engelfriet, Bottom-up and top-down tree transformations— A comparison, Math. System Theory 9 (1975) 198–231.

[10] F. Gécseg, M. Steinby, Tree Automata, Akademiai Kiado, 1984.

[11] D. Hofbauer, M. Huber, Linearizing term rewriting systems using test sets, J. Symbolic Comput. 17(1) (1994) 91–129, also in Conditional term rewriting systems (Pont-à-Mousson, 1992).

[12] G. Kucherov, On relationship between term rewriting systems and regular tree languages, in: Ron Book (Ed.), Proc. 4th Internat. Conf. on Rewriting Techniques and Applications, Como, Italy, Lecture Notes in Computer Science, Vol. 488, Springer, Berlin, April 1991, pp. 299–311.

[13] G. Kucherov, M. Tajine, Decidability of regularity and related properties of ground normal form languages, Inform. Comput. 118 (1) (1995) 91–100.

[14] S. Limet, P. Réty, E-unification by means of tree tuple synchronized grammars, in: Michel Bidoit, Max Dauchet (Eds.), TAPSOFT '97: Theory and Practice of Software Development, Lecture Notes in Computer Science, Vol. 1214, Springer, Berlin, 1997, pp. 429–440.

[15] H.A. Maurer, M. Nivat, Rational bijection of rational sets, Acta Inform. 13 (4) (1980) 365–378.

[16] M. Nivat, Transductions des langages de Chomsky, Ann. de l'Inst. Fourier 18 (1968) 339–456 (in French).

[17] S. Vágvölgyi, R. Gilleron, For a rewrite system it is decidable whether the set of irreducible ground terms is recognizable, Bull. Eur. Assoc. Theoret. Comput. Sci. 48 (1992) 197–209.